



# Non-verbal feedback, strategic signaling and non-monetary sanctioning: new experimental evidence from a public goods game

Adam Zylbersztein

## ► To cite this version:

Adam Zylbersztein. Non-verbal feedback, strategic signaling and non-monetary sanctioning: new experimental evidence from a public goods game. 2014. halshs-01098775

**HAL Id: halshs-01098775**

**<https://shs.hal.science/halshs-01098775>**

Preprint submitted on 29 Dec 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

WP 1443

**Non-verbal feedback, strategic signaling and non-monetary sanctioning: new experimental evidence from a public goods game**

Adam Zylbersztejn

December 2014

**GATE Groupe d'Analyse et de Théorie Économique Lyon-St Étienne**

93, chemin des Mouilles 69130 Ecully – France

Tel. +33 (0)4 72 86 60 60

Fax +33 (0)4 72 86 60 90

6, rue Basse des Rives 42023 Saint-Etienne cedex 02 – France

Tel. +33 (0)4 77 42 19 60

Fax. +33 (0)4 77 42 19 50

Messagerie électronique / Email : [gate@gate.cnrs.fr](mailto:gate@gate.cnrs.fr)

Téléchargement / Download : <http://www.gate.cnrs.fr> – Publications / Working Papers

# Non-verbal feedback, strategic signaling and non-monetary sanctioning: new experimental evidence from a public goods game \*

Adam Zylbersztejn<sup>†</sup>

December 16, 2014

## Abstract

Several experiments show that feedback transmission mechanisms mitigate opportunistic behavior in social dilemmas. The source of this effect, especially in a repeated interaction, nonetheless remains obscure. This study provides a novel empirical testbed for channels by which feedback may affect behavior in a repeated public goods game. One is related to strategic signaling. The other involves aversion to others' expressed disapproval. The presence of feedback is found to foster pro-social behavior. The data favour the non-monetary sanctioning explanation rather than the signaling hypothesis.

**Keywords:** Public goods game, Voluntary Contribution Mechanism, Feedback, Signaling, Non-monetary sanctioning.

**JEL Classification:** C72, D83.

---

\*This paper previously circulated under the title "Strategic signaling or emotional sanctioning? An experimental study of ex post communication in a repeated public goods game" and is a thoroughly revised version of CES WP 2013.11 and WU Wien Economics WP 161. I wish to thank Juergen Bracht, Tamás Csermely, Guillaume Fréchette, Yukihiro Funaki, Daniel Houser, Nicolas Jacquemet, Frédéric Koessler, Martin Leroch, David Masclet, Ernesto Reuben, Rupert Sausgruber, Thomas A. Stephens, Jean-Marc Tallon, Antoine Terracol, Christian Traxler, Marie-Claire Villeval and Marc Willinger for inspiring discussions. Participants at 2013 Mainz Workshop "Microfoundations of group conflict", 2013 International Meeting of the Association for Public Economic Theory in Lisbon and 2014 ADRES PhD Meeting in Paris delivered helpful comments. Maxim Frolov and Ivan Ouss provided assistance in running the experiments. This work was supported by the French National Research Agency (ANR), through the program Investissements d'Avenir (ANR-10-LABX\_93-01). Part of this research was performed within the framework of the LABEX CORTEX (ANR-11-LABX-0042) of Université de Lyon, within the program Investissements d'Avenir (ANR-11-IDEX-007) operated by the French National Research Agency (ANR).

<sup>†</sup>Université de Lyon, Lyon, F-69007, France; CNRS, GATE Lyon Saint-Etienne, 93 Chemin des Mouilles, F-69130 Ecully, France. zylbersztejn@gmail.com.

# 1 Introduction

In many economic environments, such as online trading, people interact repeatedly, remotely and anonymously. Non-verbal feedback is often used as an institutional remedy against moral hazard in these situations. For instance, users of online trading platforms (like Amazon or eBay) may rate the other traders in a positive, neutral, or negative way. Moreover, the availability of negative feedback is a key element of such rating systems. For instance, Amazon describes the rating system as follows: *we focus on negative feedback ratings as a percentage of total feedback ratings (the "Negative Feedback Rate") as an indicator of seller performance. Our very best sellers have a close to 0% Negative Feedback Rate.* In the same vein, several experimental studies (discussed in the next section) report that (i) institutions based on feedback transmission are capable of enforcing norms of pro-social behavior when individual rationality conflicts with social interest, and that (ii) the negative feedback happens to be a more powerful tool for achieving social efficiency than the positive feedback.

Although feedback transmission might be considered as a reputation building mechanism in repeated social dilemmas (Bolton, Katok, and Ockenfels, 2004), recent experiments suggest that it may affect human behavior through different channels: one involves strategic signaling (*i.e.* how the feedback recipient should act in the future), the other one stems from the idea that agents' utility may not only depend on the monetary gains generated by decisions, but also on the awareness of other people's perception (*i.e.* approval or disapproval) of these decisions.

This paper further explores the root underlying the behavioral effect of the transmission of non-verbal feedback in repeated social dilemmas, and offers a new methodology to disentangle the signaling and the sanctioning function of this mechanism. Experimental design builds on the Masclet, Noussair, Tucker, and Villeval (2003) sound intuition that under both partner matching (where the same groups of players interact many times) and stranger matching (where groups change after each interaction), feedback may work as a non-monetary sanctioning device, but only the former also allows for strategic signaling. At the same time, the experimental methodology neutralizes the undesired and potentially confounding effects of the matching protocol on subjects' behavior (Andreoni and Croson, 2008).

In line with some of the previous experiments, the presence of feedback transmitted via costless disapproval points is found to significantly foster pro-social behavior. Furthermore, the data favour the non-monetary punishment explanation rather than the strategic signaling hypothesis.

## 2 Related literature

In a seminal contribution, Masclet, Noussair, Tucker, and Villeval (2003) use a repeated public goods game, based on the voluntary contribution mechanism (VCM in short). After each round, every subject observes his group members' contributions, sends a message containing *disapproval*

*points* to each group member, and is informed about the sum of points received from others. Masclet et al. conjecture that the transmission of feedback in a repeated interaction may affect subjects' behavior in two distinct ways. First, it may serve as an information transmission device prior to the next round – for instance, signals may convey a warning that the sender will decrease his future contribution unless the receiver increases his. Second, being aware of others' opinions may operate as a non-monetary sanction/reward – for instance, people may display an aversion to being disapproved/a preference for being approved, and act less opportunistically so as to avoid/deserve it (Holländer, 1990). They argue that these two effects may be separated by contrasting subjects' behavior under partner and stranger matching. Intuitively, in the former setting the effect of such *ex post* communication may stem from both strategic information transmission and disapproval-aversion, while in the latter behavioral effects may only be a matter of subjects' aversion to disapproval. In their experiment, the presence of communication under partner matching yields significantly higher contributions than under stranger matching, which supports the strategic signaling hypothesis. In a related study, Peeters and Vorsatz (2013) use a similar experimental game and introduce treatments in which every subject may transmit an emoticon (a frowny in one condition, a smiley in the other) to each partner, and then is informed about the number of received emoticons. Comparing patterns of behavior under partner matching (where a moderate, yet statistically insignificant, effect of *ex post* communication is observed) and stranger matching (where there is virtually no effect of communication), the paper concludes that *ex post* messages mainly facilitate the exchange of strategic information before an upcoming round.

On the other hand, some experimental studies suggest that the behavioral effect of feedback may be driven by aversion to other's expressed disapproval. López-Pérez and Vorsatz (2010) report that the availability of fixed-form, post-play messages containing judgments of other participants' decisions makes subjects more cooperative in a one-shot prisoner's dilemma game. Ellingsen and Johannesson (2008) and Xiao and Houser (2009) identify the same effect on altruism in subjects playing a dictator game when a free-form, post-play communication is possible. Czap, Czap, Khachatryan, Burbach, and Lynne (2011) implement a two-stage game in which a common-pool resource is used by a group of subjects, out of which some have private incentives to produce publicly undesirable externalities. They find that the reception of a negative emotional feedback (via frownies) after the first stage reduces externalities in the second stage, while providing positive feedback (via smilies) does the opposite. Dugar (2010) studies a repeated weakest-link coordination game and finds that the introduction of disapproval points helps groups of players converge towards the Pareto-superior Nash equilibrium, while approval points bring no improvement in this respect, inhibiting a fast convergence to the Pareto-inferior Nash equilibrium. Dugar (2013) finds that the availability of disapproval points generates higher contributions than approval points in a fixed-group VCM game based on Masclet, Noussair, Tucker, and Villeval (2003), and that combining *both* kinds of points brings a further improvement with this respect. Dugar attributes the effects

observed in these two experiments to an asymmetrical sensitivity to approval and disapproval.

### 3 Empirical strategy

The experimental methodology introduced by Masclet, Noussair, Tucker, and Villeval (2003) and subsequently used by Peeters and Vorsatz (2013) has its strengths and weaknesses for separating the signaling dimension and the sanctioning dimension of non-verbal feedback in a repeated interaction. Its advantage is that it identifies environments where both of these phenomena either can or cannot co-exist. The main disadvantage is that these environments may affect behavior simultaneously to the content of messages, and that these behavioral effects are highly unpredictable.

As a simple illustration of the above problem, consider a situation where only two factors affect contributions: received feedback and matching protocol. Since both effects arise simultaneously, future contributions are affected not only by past communication, but also by the matching protocol. The two effects should not be considered orthogonal: assuming that the content of feedback is correlated with own and others' past behavior (which is the case in most studies of this type of communication, including this one), future messages depend on past contributions, so that the very nature of feedback is also influenced by the matching protocol. Consequently, the matching protocol effect is a confound that aggravates the isolation of the actual relationship between communication and contributions in a repeated game. Thus, the most general solution to this issue is an experimental design in which actions are uncorrelated with the matching scheme.<sup>1</sup>

The central methodological challenge for this study is to create an experimental environment that allows for different features of feedback, at the same time ruling out the unobserved effect of the matching protocol on decisions. To this end, I introduce a novel uniform matching protocol where a random process decides whether groups are maintained or broken up. In each round of a repeated game, subjects decide upon their level on contribution before learning whether their group prevails until next round. This design neutralizes the unwanted forward-looking effects of

---

<sup>1</sup>Masclet, Noussair, Tucker, and Villeval (2003) offer a way to control for the matching protocol effects with a difference-in-difference approach: in every session subjects play a sequence of rounds without communication, which is followed by an analogous sequence with communication, and then another sequence without communication, holding the matching protocol constant. They argue that since subjects' behavior is similar in the initial sequence under both matching schemes, thus the differences observed in the second stage are unlikely to stem from the matching protocol effects. However, it should be also noticed that a conclusive comparison in the second stage is only possible in the absence of a systematic partner-stranger difference in the first stage. Unfortunately, scenarios where observations are neutral to matching protocol effects are far from being a regularity in lab experiments on public goods games. In a related study, Peeters and Vorsatz (2013) use a classical between-subject design and observe systematic matching protocol effects: in each round of every treatment, partner matching induces higher contributions than stranger matching (see the working paper version of their study, Peeters and Vorsatz (2009)). Moreover, the matching protocol effects are absent in only 4 out of 13 experiments overviewed by Andreoni and Croson (2008). In 5 cases partner matching brings higher contribution than stranger matching, in the remaining 4 cases the opposite occurs.

different matching protocols on contributions, while controlling for the backward-looking factors (as discussed in the next section). The transmission of feedback, in turn, only takes place after the fate of groups is known to subjects. Consequently, the effect of communication can be captured in two different strategic contexts: when groups prevail from one period to another, and when they change between rounds. In line with Masclet et al.’s original argument, strategic information transmission is only possible in the former case, while non-monetary punishment may occur in both cases.

Following several previous studies (Masclet, Noussair, Tucker, and Villeval, 2003; Noussair and Tucker, 2005; Dugar, 2010, 2013), I implement a communication protocol based on numerical *ex post* messages framed as disapproval points. In order to assure the interpretability of messages, the VCM game is played by groups of two subjects, so in each round every participant learns about other group member’s contribution, sends a message and receives one in return. Consequently, messages may be easily matched to individual actions, which creates an environment that (i) facilitates agents’ comprehension of non-verbal content, and (ii) allows the experimenter to trace the relationship between individual messages and individual contributions.<sup>2</sup> Moreover, it is of common knowledge that no group ever re-appears after having been dissolved, which is aimed at ruling out the possibility of sending future-oriented signals between subjects who are about to cease interacting.

### 3.1 Experimental game and conditions

**Experimental game.** Pairs of subjects play the following version of the VCM game. Each player holds an initial endowment of 15 Experimental Currency Units (ECU), and may contribute any part of it to the common pool.<sup>3</sup> Decisions are made simultaneously and the amount accumulated in the common pool is then multiplied by 1.5 and re-transferred to group members in equal parts. Thus, the gain of player  $i$  who contributes  $N_i$  and interacts with player  $j$  who contributes  $M_j$  equals:

$$Gain_i = 15 - 0.25 \times N_i + 0.75 \times M_j \quad (1)$$

Although the social welfare is maximized when each player contributes his entire endowment, the dominant strategy is to contribute nothing.<sup>4</sup> Therefore, the standard game theory predicts

---

<sup>2</sup>This two-person design, motivated by the clarity of communication, comes as a departure from the four-person design used in the related studies by Masclet, Noussair, Tucker, and Villeval (2003) and Peeters and Vorsatz (2013). However, experimentals by Isaac and Walker (1988) and Isaac, Walker, and Williams (1994) suggest that the group size does not *per se* affect contributions in VCM.

<sup>3</sup>In the lab implementation, contributions may only take integer values between 0 ECU and 15 ECU.

<sup>4</sup>To avoid framing effects, instructions use neutral phrasing: I use expressions such as *players* and *group members* rather than *partners*, and *contributions* are never related to *cooperation*. See Rege and Telle (2004) for evidence on framing effects in public goods experiments.



that in the unique Nash equilibrium all players contribute nothing.

**Baseline condition.** In the baseline condition (BC), the game is repeated as follows. In each occurrence, subjects *(i)* decide upon their contribution to the common pool, *(ii)* learn whether their current group survives until the next round, and finally *(iii)* observe the other group member’s contribution, as well as their own gain. Subjects are informed that in each round their groups survive with a 50% chance, that this process is entirely random, and that every change is permanent – groups that disappear cannot re-appear in the future. In all rounds following round 1, an announcement prior to the initial stage reminds subjects whether their group has changed with respect to the previous period. The important issue of the asymmetry of information about players’ past behavior between maintained and newly formed groups is addressed in the following way. Before stage *(i)*, members of newly formed pairs are informed about the contribution recently made by their current group member in his former group.<sup>5</sup> Although the extent of historical information may vary between new and old pairs due to the randomness of group re-matching, this design maintains the *minimum* level of knowledge about group members’ past.<sup>6</sup>

**Communication condition.** The communication condition (CC) encompasses the three stages forming BC, as well as the current-group-status reminder. In addition, in stage *(iv)* subjects are asked to express their opinion about a group member’s decision by assigning a certain number of points (between 0 and 10) to him. Experimental instructions state that *a high number of points expresses disapproval: 10 points correspond to the strongest disapproval, while 0 points correspond to the weakest disapproval*, and that *assigned points do not affect either participant’s gains for the experiment*. Then, each subject is informed about the number of points he received from the other group member. If groups change between periods  $t - 1$  and  $t$ , prior to stage *(i)* subjects not only learn about the decision taken by their current group member in  $t - 1$  (like in BC), but also about the number of points he received.

### 3.2 Experimental procedures

The experiment involves a total of 12 sessions (6 for BC and 6 for CC), each comprising 8 subjects. I use a round-robin matching protocol (ensuring that every two subjects have the opportunity to interact during the experiment) and the random group survival rule outlined above. Consequently, the structure of group matching and the number of rounds may differ between sessions. In order to control for the effects of these variations, I use six independent, randomly generated matching

---

<sup>5</sup>Instructions (translated from French to English) are provided as a supplementary material.

<sup>6</sup>Another advantage of the present design is related to reputation-building. One may argue that this design may involve reputation-building. However, even if subjects indeed try to establish a reputation, this incentive remains constant throughout the entire experiment: across different kinds of pairs – newly established and preserved, as well as under different communication conditions. Yet, this symmetry is another important refinement with respect to the usual partner-stranger comparison in which reputation-building may unobservedly arise in the former scheme, but not in the latter.

sequences (henceforth labeled *Game 1*, ..., *Game 6*) and run two separate sessions for each of them: one implementing BC and one implementing CC.<sup>7</sup> Subjects are informed that the game contains between 10 and 16 rounds and that its length is determined randomly. In practice, sessions contain between 11 and 15 rounds, and the pairs of subjects interact for up to five consecutive rounds.<sup>8</sup>

At the beginning of each session, participants are randomly assigned to their computers and asked to fill in a small personal questionnaire containing basic questions about their age, gender, education, *etc.* The written instructions are then read aloud. Before starting, subjects are also asked to fill in a quiz assessing their understanding of the game they are about to play. Once the quiz and all remaining questions are answered, the experiment begins.

Once all pairs complete a round of the game, subjects are informed that either a new round will start, or the experiment will end. In the latter case, a single round is randomly drawn and each participant receives the amount in EUR corresponding to his gains in that round (converted from ECU to EUR using an exchange rate  $1 \text{ ECU} = 0.40 \text{ EUR}$ ), plus a show-up fee equal to 5 EUR.

All sessions took place in the lab of the University of Paris 1 (LEEP) in July 2012. The recruitment of subjects was carried out via LEEP database among individuals who have successfully completed the registration process on laboratory's website.<sup>9</sup> Among 96 participants, 51 are male and 45 are female. 63 participant are students, of which 67% might have some background in game theory due to their field of study.<sup>10</sup> 82 subjects had taken part in experiments organized in LEEP in the past. Participants' average age is roughly 25. No subject participated in more than one experimental session. Each session lasted about 45 minutes with average earning of 12.20 Euros.

## 4 Results

This section establishes the principal results coming from individual- and group-level data analysis. The main findings are as follows. First, the transmission of feedback involving disapproval points foster the production of public goods. Second, the effects of *ex post* communication on group productivity and the within-group structure of contributions do not vary between the cases where messages may or may not have a strategic (*i.e.* future-oriented) function.

---

<sup>7</sup>The exact structure of these sequences can be found in the supplementary material.

<sup>8</sup>In the remainder of the paper, I use the terms *pair* and *group* interchangeably.

<sup>9</sup>The recruitment uses ORSEE Greiner (2004); the experiment is computerized through a software developed under REGATE Zeiliger (2000).

<sup>10</sup>Disciplines such as economics, engineering, management, political science, psychology, mathematics applied in social science, mathematics, computer science, sociology, biology.

Table 1: Average contributions by treatment and experimental game

Conditions	Game 1	Game 2	Game 3	Game 4	Game 5	Game 6	Average	$p$
Average contributions in round 1								
Baseline	7.625	6.750	8.125	6.375	5.750	3.875	6.417	0.210
Evaluation	6.250	8.250	6.500	8.375	7.875	8.750	7.667	
Overall average contributions								
Baseline	5.000	2.900	4.420	3.083	6.010	2.846	3.942	0.046
Evaluation	4.083	7.783	6.045	6.325	9.135	8.029	6.938	
Session details								
Number of subjects	8	8	8	8	8	8		
Number of rounds	12	15	11	15	12	13		

**Note.** Columns 1-6 present average contributions in each experimental game, using data from round 1 (upper part) and all rounds (middle part). The last two columns summarize these results and provide non-parametric tests for the significance of the effect of treatment on contributions: the Wilcoxon-Mann-Whitney rank-sum test using individual observations in round 1, and Wilcoxon signed-rank test using game-level matched averages for the aggregate data (in order to account for the potential sequence-specific effects). The lower part of the table contains additional information on the number of subjects and the length of each experimental game.

#### 4.1 Individual behavior

Table 1 presents subjects' average contributions as a function of the structure of the experimental game and the communication condition. In five experimental games out of six, the presence of non-verbal feedback increases the average contribution. This shift in behavior is significant at the 5% level according to a Wilcoxon signed-rank test.<sup>11, 12</sup> Then, the behavior of inexperienced players is neutral to the treatment variable: a Wilcoxon-Mann-Whitney rank-sum test using individual observations from round 1 does not reject the hypothesis that the distributions of subjects' choices are the same in both experimental environments ( $p=0.210$ ).<sup>13</sup>

**Result 1.** The average contribution in the communication condition is significantly higher than in the baseline condition. However, this effect does not occur in the initial round.

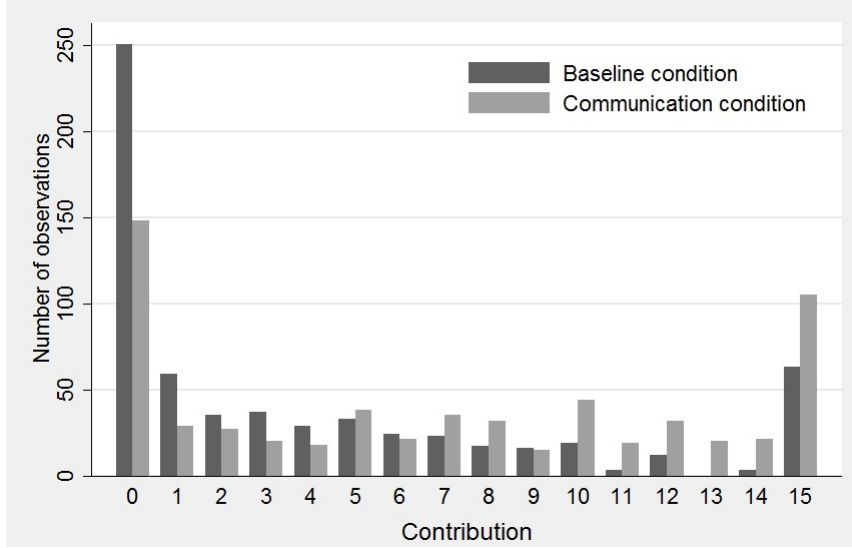
To further investigate the behavioral transition caused by communication, Figure 1 compares the distributions of contributed amounts in both experimental conditions. In BC, over 40% (251 out of 624) of decisions are zero contribution, as compared with less than 25% (148/624) in CC. Moreover, contributions between 1 ECU and 4 ECU are more frequent in the former than in the latter. The relationship between the two conditions becomes unstable until the threshold level of 10 ECU, above which all values appear substantially more often in CC than in BC, including the case in which the entire endowment is transferred to the common pool (105 times in CC

<sup>11</sup>For each of the six games, this test matches the average contributions observed in BC and in CC, and therefore accounts for the effects related to different game structures.

<sup>12</sup>All  $p$ -values used in statistical analysis correspond to two-sided tests.

<sup>13</sup>This may suggest that the effect of communication via disapproval points involves learning of conventions and arises via procedural experience. An identical phenomenon is reported by Dugar (2010, 2013).

Figure 1: Distribution of contributions across treatments



**Note.** For each experimental treatment, data contain 624 observations from 6 experimental sessions. Contributions are given in ECU.

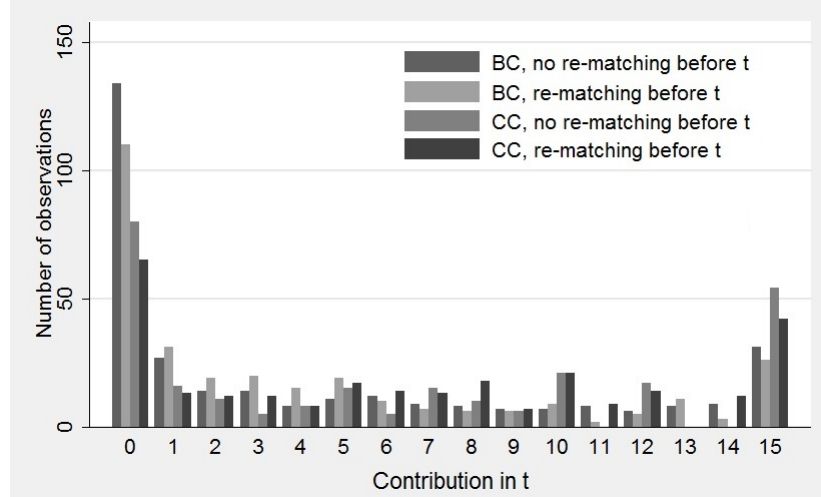
against 63 times in BT). In addition, Figure 2 suggests that this phenomenon is robust to the way pairs are matched prior to interacting. Within each experimental condition, one observes similar distributions of contributions in newly formed and retained pairs. On the other hand, the key difference between both conditions observed in Figure 1 prevails: in the presence of feedback the frequency of low contributions is substantially lower than without feedback, while the opposite holds for high contributions.<sup>14</sup>

## 4.2 Group behavior

In order to uncover the cause of the transition in subjects' decision-making, let us now turn to exploring the behavior of groups in which this decision-making (either alone or couple with communication) take place. This analysis is aimed at providing statistical tests of the signaling and non-monetary sanctioning hypotheses via parametric regression models. The core instruments are embedded in the  $2 \times 2$  design of the experiment: the first dimension being the presence of random  $1_{Re-matching_{t-1}^t}$  (set to 1 if re-matching occurs, 0 otherwise) of pairs prior to round  $t$ , while the second – the presence of *ex post* communication treatment ( $1_{CC} = 1$  for communication

<sup>14</sup>For instance, zero contributions are observed 134 times in BC in the absence of prior re-matching of pairs, and 110 times had re-matching occurred. Analogous figures in CC are 80 and 65. Moving to the opposite extreme, contributions of 15 occur 31, 26, 54 and 42 times in the four respective cases.

Figure 2: Contributions, communication and re-matching



**Note.** For each combination of experimental condition (BC or CC) and pair status (re-matching occurs before period  $t$  or not), data contain 288 observations from 6 experimental sessions. Contributions are given in ECU.

condition, 0 otherwise). The statistical analysis focuses on two dependent variables:

*Contributions gathered by players  $i$  and  $j$  interacting at time  $t$  ( $c_t^i + c_t^j$ ).* The signaling role of *ex post* communication may be confirmed by the significance of  $1_{CC}$  in maintained pairs, and a significant difference between the contributions accumulated in maintained and re-matched pairs when  $1_{CC} = 1$  – both of which indicate that the effect of feedback on group productivity depends on the continuity of interaction. On the other hand, the absence of a significant difference between contributions in maintained and re-matched pairs when  $1_{CC} = 1$  points towards the importance of the non-monetary sanctioning hypothesis. These mechanisms are tested in Model 1.

*The absolute difference between the contributions of players  $i$  and  $j$  who form a pair at time  $t$  ( $|c_t^i - c_t^j|$ ).* Model 2 provides a complementary test for the signaling hypothesis. If non-verbal feedback operates as an efficient signaling mechanism, then it should facilitate players' mutual strategic adaptation resulting in more equilibrated contributions within pairs. Consequently, the contributions in maintained groups playing under the communication condition should involve less intra-group inequality than elsewhere.

Table 2 summarizes the estimates from both models.<sup>15</sup> Model 1 reveals that the presence of communication increases the provision of public good in both retained and re-matched groups

<sup>15</sup>Supplementary material provides an additional specification – double-censored tobit – as an alternative to the OLS estimates reported in Tables 2-4. The OLS specification is more conservative regarding the measurement of treatment effects, and thus chosen for this analysis. However, all the effects discussed in this section have been captured by both models.

Table 2: Determinants of group behavior

Dep. variable:	Model 1 $c_t^i + c_t^j$	Model 2 $ c_t^i - c_t^j $
	$\beta$	$\beta$
<i>Intercept</i> ( $\beta_0$ )	8.032	4.422*
$1_{Re-matching_{t-1}^i}$ ( $\beta_1$ )	0.753	0.928*
$1_{CC}$ ( $\beta_2$ )	6.375*	0.838
$1_{CC} \times 1_{Re-matching_{t-1}^i}$ ( $\beta_3$ )	-0.188	-0.785
Additional controls:		
<i>Round</i>		
3	-1.627	1.580
4	-0.938	0.938
5	-2.803**	0.551
6	-2.241**	0.551
7	-2.960**	0.080
8	-3.860**	-0.839
9	-3.741***	0.592
10	-5.205***	-0.976
11	-4.553**	0.488
12	-4.317***	0.704
13	-4.587**	0.687
14	-3.550**	1.666
15	-4.334**	0.630
<i>Game :</i>		
2	2.067	-0.616
3	1.228	-1.408
4	0.719	-1.268
5	6.636	-0.386
6	2.249	-1.263
<i>N</i>	576	
<i>R</i> <sup>2</sup>	0.208	0.043

**Note.** OLS regressions using group-level observations. Dependent variable in Model 1 is the sum of contributions of players  $i$  and  $j$  forming a group at time  $t$  ( $c_t^i + c_t^j$ ), whereas Model 2 uses the absolute difference of contributions within this group ( $|c_t^i - c_t^j|$ ). Columns contain model's coefficients ( $\beta$ ) and corresponding  $p$ -values, with \*/\*\*/\*\* indicating significance at the 10%/5%/1% level. Explanatory variables include a dummy indicating the occurrence of group re-matching prior to ( $1_{Re-matching_{t-1}^i}$ ) and another dummy indicating whether the game involves *ex post* communication ( $1_{CC}$ ), as well as their interaction. Models also control for the round and matching-sequence effects, with *Round 2* and *Game 1* being set as reference points and omitted in regressions. Data include observation from round 2 onwards. Standard errors are clustered at the session level (12 clusters) and are corrected using leave-one-out jackknife.

Table 3: Contributions and historical information in newly formed pairs

<i>Intercept</i> ( $\theta_0$ )	4.133***	7.200***	4.304**
<i>Group member's:</i>			
<i>Contribution</i> $_{t-1}$ ( $\theta_1$ )	0.228***	–	0.291**
<i>Received_points</i> $_{t-1}$ ( $\theta_2$ )	–	-0.080	0.158
<i>N</i>	288		
<i>R</i> <sup>2</sup>	0.053	0.003	0.060

**Note.** OLS regressions of player's contribution in  $t$  on the historical information about the current group member: his contribution in  $t - 1$  and the number of points he received in  $t - 1$ . Columns contain model's coefficients ( $\theta$ ), with \*/\*\*/\*\* indicating significance at the 10%/5%/1% level. Data come from newly matched pairs. Standard errors are estimated using observations clustered at the session level (6 clusters in total) and corrected using leave-one-out jackknife.

( $H_0 : \beta_2 = 0, p = 0.058$ ,  $H_0 : \beta_2 + \beta_3 = 0, p = 0.023$ , respectively). Moreover, the group behavior in the communication condition does not depend on the previous occurrence of re-matching ( $H_0 : \beta_1 + \beta_3 = 0, p = 0.497$ ). Altogether, this evidence favours the non-monetary sanctioning explanation over the signaling hypothesis: the impact of communication on the group behavior is significant, but does not depend on players' ability to transmit strategic signals.

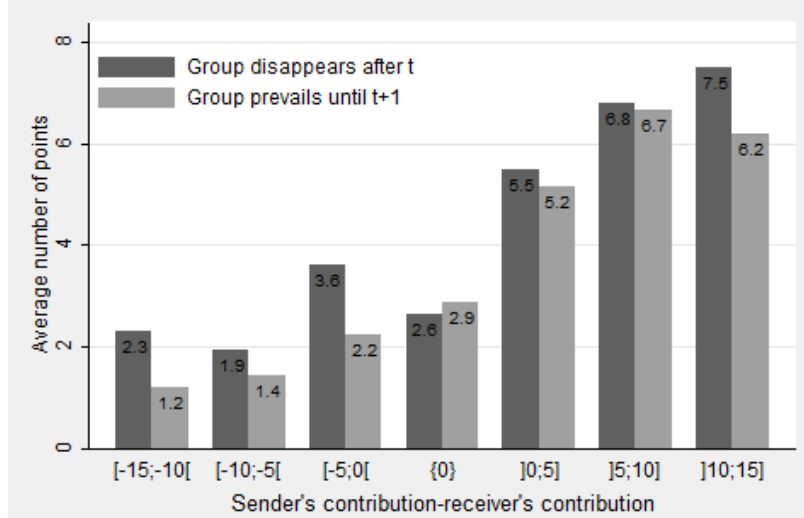
Model 2 provides evidence in the same vein: the within-group inequality in maintained pairs does not differ systematically between BC and CC. One cannot reject the joint hypothesis that the level of inequality in maintained pairs is not affected by the presence of *ex post* communication, and that the level of inequality in the communication condition does not depend on the nature of groups' re-matching ( $H_0 : \beta_2 = 0 \cap \beta_1 + \beta_3 = 0, p = 0.464$ ). On the other hand, the contributions within newly formed groups in the baseline condition tend to be more divergent than in retained groups ( $H_0 : \beta_1 = 0, p = 0.099$ ), which suggests that the experience acquired over the course of a continuous interaction improves players' capacity of strategic adaptation, and that this effect is superseded by the *ex post* exchange of messages.

**Result 2.** The effect of communication on group productivity and players' strategic adaptation within groups does not depend on the availability of future-oriented signals.

Finally, both regressions also reveal interesting time trends: a gradual decay in the production of public goods in Model 1<sup>16</sup>, and no systematic variations in within-group inequalities in Model 2.

<sup>16</sup>This observation echoes previous experiments on public goods with and without feedback mechanisms. See, for instance, Masclet, Noussair, Tucker, and Villeval (2003); Dugar (2013); Peeters and Vorsatz (2013).

Figure 3: Sent disapproval points and the relative size of contribution



**Note.** Data come from 6 sessions and contain 288 observations for each type of group. Contributions are given in ECU.

### 4.3 Does re-matching preclude strategic signaling? A simple test of the altruistic signaling hypothesis

Note, the underlying assumption of the above analysis is that the transmission of strategic signals is only possible in continuous interactions. However, since players in newly established pairs observe the degree of disapproval their group members received in their old pairs, one may also consider an alternative scenario – altruistic signaling – where disapproval points are utilized to label players who enter into a new interaction. More precisely, subsequent members of each subject's pair may transmit cues about this person to their successors and adapt their decisions according to the cues transmitted by their predecessors.

To test this hypothesis, the data from newly matched pairs are used to regress the individual contributions on the historical information held by each player about his current group member: his past contribution and the number of disapproval points he received in the previous period. The estimates are presented in Table 3. This simple empirical test suggests that re-matching precludes the strategic use of these points: while group members' past behavior matters for subjects' choices, disapproval points the former received in their previous groups do not.

### 4.4 Determinants of expressed disapproval

This section addresses two central questions related to the formation of *ex post* messages: (i) what is the relationship between the content of feedback and underlying contributions, and (ii) does



this relationship vary between retained and re-matched groups? The answer to the first question is provided by Figure 3 which presents the average number of points sent by subject to his group member as a function of his deviation from other group member’s contribution, and the outcome of the random draw that determines group’s future. Under both matching scenarios, the overall pattern seems rather coherent: the number of sent disapproval points grows as the difference between sender’s and receiver’s contributions grows.

Answering the second question requires a more formal insight into the process of message formation. For this sake, I estimate a linear model representing the number of points sent by a player in round  $t$  as a function of his *Group\_member’s\_contribution<sub>t</sub>*, as well as its relative size within his group – *i.e.* the magnitude of *Sender’s\_positive\_deviation<sub>t</sub>* or *Sender’s\_absolute\_negative\_deviation<sub>t</sub>* with respect to other group member’s decision. The effect of prospective re-matching is captured by the variable  $1_{Re-matching_t^{t+1}}$  (like before, set to 1 if re-matching is about to occur between  $t$  and  $t + 1$ , a 0 otherwise) as well as its interactions with the above variables.

In the absence of re-matching prior to the upcoming interaction, players send more disapproval points the less their group members contribute ( $H_0 : \gamma_1 = 0, p = 0.002$ ), and the more they contribute themselves relative to their group members ( $H_0 : \gamma_2 = 0, p = 0.010$ ). The magnitude of own *negative* deviation (in absolute terms), in turn, does not play a significant role in the process of point attribution ( $H_0 : \gamma_3 = 0, p = 0.329$ ). Importantly, these effects do not change due to re-matching. One may not reject the joint hypothesis that the prospect of group re-matching does not change the relationship between one’s relative contribution and the number of disapproval points he sends ( $H_0 : \gamma_5 = 0 \cap \gamma_6 = 0 \cap \gamma_7 = 0, p = 0.356$ ).

**Result 3.** *Ex post* disapproval points are assigned in an informative and coherent manner: low contributors receive more disapproval points than high contributors. This pattern of point attribution does not vary significantly between pairs that are about to cease and continue interacting.

## 5 Summary and discussion

Recent developments in economic research suggest that communication may mitigate selfishness in social dilemmas, and the source of this phenomenon is often explained as an emotional reaction communication evokes in humans. For instance, Ellingsen and Johannesson (2004) and Vanberg (2008) argue that *ex ante* communication may reduce the scope of opportunistic behavior due to agents’ *aversion to lying*, while Charness and Dufwenberg (2006) relate this transition to *guilt aversion* due to which agents experience disutility from letting down others’ expectations.

The present study provides new evidence on the behavioral effect of communication in social dilemmas First, echoing the previous findings by Masclet, Noussair, Tucker, and Villeval (2003) and Dugar (2013) (and notwithstanding the findings by Peeters and Vorsatz (2013)), the paper reports that the availability of a feedback mechanism involving costless disapproval points substan-

Table 4: Sent disapproval points and the relative size of contribution

<i>Intercept</i> ( $\gamma_0$ )	6.227***
<i>Group_member's_contribution</i> <sub><i>t</i></sub> ( $\gamma_1$ )	-0.380***
<i>Sender's_positive_deviation</i> <sub><i>t</i></sub> ( $\gamma_2$ )	0.167***
<i>Sender's_absolute_negative_deviation</i> <sub><i>t</i></sub> ( $\gamma_3$ )	0.075
$1_{Re-matching_t^{t+1}}$ ( $\gamma_4$ )	-0.536
$1_{Re-matching_t^{t+1}} \times Group\_members's\_contribution_t$ ( $\gamma_5$ )	0.047
$1_{Re-matching_t^{t+1}} \times Sender's\_positive\_deviation_t$ ( $\gamma_6$ )	0.010
$1_{Re-matching_t^{t+1}} \times Sender's\_absolute\_negative\_deviation_t$ ( $\gamma_7$ )	-0.138
<i>N</i>	576
<i>R</i> <sup>2</sup>	0.361

**Note.** OLS regressions of the number of points sent in  $t$  on variables indicating the level of *Group\_member's\_contribution*<sub>*t*</sub>, the value of *Sender's\_positive\_deviation*<sub>*t*</sub> or *Sender's\_absolute\_negative\_deviation*<sub>*t*</sub>, a dummy indicating the occurrence of  $1_{Re-matching_t^{t+1}}$  as well as its interactions with the three above variables. \*/\*\*/\*\* indicate coefficients' significance at the 10%/5%/1% level. Standard errors are estimated using observations clustered at the session level (6 clusters in total) and corrected using leave-one-out jackknife.

tially reduces opportunistic behavior in a repeated VCM game. Second, it offers a new angle for understanding this important behavioral process and to test two potential (and non-excludable) factors behind the behavioral effect of feedback: first, the preference for approval/aversion to disapproval, understood as non-monetary utility or disutility drawn from others' opinions about own behavior; second, the transmission of strategic signals (such as threats) linked to future interactions. It builds on the intuition that under both a repeated interaction within fixed groups of players (partner protocol) and interactions with constantly changing groups (stranger protocol) *ex post* communication may be used for non-monetary sanctioning or rewarding, but only the former also allows for future-oriented strategic signaling. While retaining these different features of communication, this new design neutralizes the undesired effects of matching protocols on contributions.

The experimental literature lacks unequivocal evidence on the root underlying the behavioral effect of *ex post* communication in repeated social dilemmas. For instance, Masclet, Noussair, Tucker, and Villeval (2003) suggest that both of these mechanisms may influence subjects' behavior, Peeters and Vorsatz (2013) highlight the role of strategic information transmission, while Dugar (2013) points towards the importance of the utility or disutility induced by others' judgments about own actions. Data from this experiment do not support the signaling hypothesis, favouring the non-monetary punishment explanation instead. First, the provision of public good is invariant across groups that may exchange future-oriented messages and those whose messages cannot be future-oriented. Although communication increases the production of public good irrespective of the continuity of group's interaction, it does not improve players' strategic adaptation

(measured as the within-group inequality of contributions), even if interactions extend throughout multiple periods. Second, the relationship between one's relative contribution and the resulting level of others' disapproval is coherent (low contributors face stronger disapproval than high contributors) and does not vary as a function of the continuity of interaction.

In a broader perspective, this experiment also contributes to furthering our understanding of the determinants of the behavioral response to communication. Charness and Dufwenberg (2010) discuss two implementations of pre-play, cheap-talk, natural language communication in a trust game: an unstructured (free-form) protocol where (almost) any content may be transmitted between players, and a fixed-form protocol which only allows for single-phrase, standardized "bare" promises. Their conclusion is that only unrestricted communication may induce emotional reaction from subjects (which, in their study, is explained as an emotional cost caused by lying-aversion). This conjecture finds support in some experiments contrasting non-verbal and verbal *ex ante* communication. Wilson and Sell (1997) find no effect of numerical pre-play communication in an experimental VCM game. Using similar setup, Bochet, Page, and Putterman (2006) again observe no effect of numerical pre-play announcements, while verbal pre-play communication is found to foster cooperativeness. Finally, Ben-Ner, Putterman, and Ren (2011) find that both mechanisms improve the efficiency of subjects' behavior in a trust game, with verbal communication systematically outperforming numerical communication. *Ex post* communication based on natural language was also found to affect subjects' emotions in social dilemmas (Ellingsen and Johannesson, 2008; Xiao and Houser, 2009; López-Pérez and Vorsatz, 2010). On the other hand, Dugar (2010, 2013) suggests that non-verbal *ex post* communication may also refer to emotions. An important result pinned down by the present experiment differs at face value from Charness and Dufwenberg's insight, echoing Dugar's conjecture. Even a structured, artificial and wordless method of expressing own judgment – such as the utilization of evaluation points – creates an environment in which aversion to others' disapproval may refrain opportunistic behavior in social dilemmas. A more systematic comparison of verbal and non-verbal feedback mechanisms could certainly constitute a desirable avenue for future research.

## References

- ANDREONI, J., AND R. CROSON (2008): *Partners versus Strangers: Random Rematching in Public Goods Experiments* vol. 1 of *Handbook of Experimental Economics Results*, chap. 82, pp. 776–783. Elsevier.
- BEN-NER, A., L. PUTTERMAN, AND T. REN (2011): "Lavish returns on cheap talk: Two-way communication in trust games," *The Journal of Socio-Economics*, 40(1), 1 – 13.
- BOCHET, O., T. PAGE, AND L. PUTTERMAN (2006): "Communication and punishment in voluntary contribution experiments," *Journal of Economic Behavior & Organization*, 60(1), 11–26.

- BOLTON, G. E., E. KATOK, AND A. OCKENFELS (2004): “How Effective Are Electronic Reputation Mechanisms? An Experimental Investigation,” *Management Science*, 50(11), 1587–1602.
- CHARNESS, G., AND M. DUFWENBERG (2006): “Promises and Partnership,” *Econometrica*, 74(6), 1579–1601.
- (2010): “Bare promises: An experiment,” *Economics Letters*, 107(2), 281–283.
- CZAP, H. J., N. V. CZAP, M. KHACHATURYAN, M. E. BURBACH, AND G. D. LYNNE (2011): “Smiley or Frowney: The effect of emotions and framing in a downstream water pollution game,” Discussion paper.
- DUGAR, S. (2010): “Nonmonetary sanctions and rewards in an experimental coordination game,” *Journal of Economic Behavior & Organization*, 73(3), 377–386.
- (2013): “Non-monetary incentives and opportunistic behavior: evidence from a laboratory public good game,” *Economic Inquiry*, 51(2), 1374–1388.
- ELLINGSEN, T., AND M. JOHANNESSON (2004): “Promises, Threats and Fairness,” *Economic Journal*, 114(495), 397–420.
- (2008): “Anticipated verbal feedback induces altruistic behavior,” *Evolution and Human Behavior*, 29(2), 100–105.
- GREINER, B. (2004): “An Online Recruitment System for Economic Experiments.,” in *Forschung und wissenschaftliches Rechnen 2003. GWDG Bericht 63*, ed. by K. Kremer, and V. Macho, pp. 79–93. Ges. für Wiss. Datenverarbeitung, Göttingen.
- HOLLÄNDER, H. (1990): “A Social Exchange Approach to Voluntary Cooperation,” *The American Economic Review*, 80(5), pp. 1157–1167.
- ISAAC, R. M., AND J. M. WALKER (1988): “Group Size Effects in Public Goods Provision: The Voluntary Contributions Mechanism,” *The Quarterly Journal of Economics*, 103(1), pp. 179–199.
- ISAAC, R. M., J. M. WALKER, AND A. W. WILLIAMS (1994): “Group size and the voluntary provision of public goods : Experimental evidence utilizing large groups,” *Journal of Public Economics*, 54(1), 1–36.
- LÓPEZ-PÉREZ, R., AND M. VORSATZ (2010): “On approval and disapproval: Theory and experiments,” *Journal of Economic Psychology*, 31(4), 527–541.
- MASCLET, D., C. NOUSSAIR, S. TUCKER, AND M.-C. VILLEVAL (2003): “Monetary and Nonmonetary Punishment in the Voluntary Contributions Mechanism,” *American Economic Review*, 93(1), 366–380.
- NOUSSAIR, C., AND S. TUCKER (2005): “Combining Monetary and Social Sanctions to Promote Cooperation,” *Economic Inquiry*, 43(3), 649–660.
- PEETERS, R., AND M. VORSATZ (2009): “Immaterial rewards and sanctions in a voluntary contribution experiment,” Discussion paper.

- (2013): “Immaterial Rewards And Sanctions In A Voluntary Contribution Experiment,” *Economic Inquiry*, 51(2), 1442–1456.
- REGE, M., AND K. TELLE (2004): “The impact of social approval and framing on cooperation in public good situations,” *Journal of Public Economics*, 88(7-8), 1625–1644.
- VANBERG, C. (2008): “Why Do People Keep Their Promises? An Experimental Test of Two Explanations,” *Econometrica*, 76(6), 1467–1480.
- WILSON, R. K., AND J. SELL (1997): “"Liar, Liar...": Cheap Talk and Reputation in Repeated Public Goods Settings,” *The Journal of Conflict Resolution*, 41(5), pp. 695–717.
- XIAO, E., AND D. HOUSER (2009): “Avoiding the sharp tongue: Anticipated written messages promote fair economic exchange,” *Journal of Economic Psychology*, 30(3), 393–404.
- ZEILIGER, R. (2000): “A presentation of Regate, Internet based Software for Experimental Economics,” <http://www.gate.cnrs.fr/~zeiliger/regate/RegateIntro.ppt>, *GATE*.

# Supplementary material

## Written instructions

*Author's note: BC (EC) at the beginning of a paragraph indicates that this part is specific to baseline condition (evaluation condition)*

You are about to take part in an experiment in which you can earn money. Your gains will depend on your decisions, as well as on the decisions made by other participants.

Before starting, we would like to ask you to answer a few standard questions (concerning your age, education, profession, ...) that will help us to get to know you better. **This information, as well as the amount of your gains in this experiment, will remain strictly confidential and anonymous.**

Please, fill in the questionnaire using the interface on your computer screen, which is divided into three parts:

- In the *top* section, you will find information that might help you in making decisions.
- In the *middle* section, you will submit your decisions by clicking on a relevant button.
- In the *bottom* section, you will see all your decisions and gains from previous rounds of the experiment.

**Thank you.**

## THE EXPERIMENT

The experiment consists of several rounds. The total number of rounds is random and might vary between 10 and 16. In each round, participants are divided into groups of two. In each round (more precisely, in Stage 2 described below) a random draw decides that:

- either your group will not change in the next round;
- or that your group will change in the next round. In order to form your new group, another participant will be chosen at random among participant who have never been part of your group before.

**Both events are equally probable – each occurs with a 50% chance.**

In every round, each participant's gain is determined as follows. At the beginning of round, every person possesses the initial endowment of 15 ECU (Experimental Currency Unit).

Members of each group may create a common pool. Each participant freely chooses his level of contribution to the pool that may range between 0 ECU and 15 ECU. The total amount gathered in the pool is then multiplied by 1.5 and divided equally among group members.

For instance, if you are participant  $i$  who contributed  $N_i$ , and the other member of your group, participant  $j$ , contributed  $M_j$ , then your gain ( $Gain_i$ ) equals:

$$Gain_i = 15 - 0.25 \times N_i + 0.75 \times M_j \quad (2)$$

The Table provided below presents your gain in ECU in a given period as a function of your level of contribution and the other group member's level of contribution.

Table 5: Your gain in ECU in a given period as a function of your decision and your group member's decision

		Your group member's level of contribution															
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Your level of contribution	0	15	15.75	16.5	17.25	18	18.75	19.5	20.25	21	21.75	22.5	23.25	24	24.75	25.5	26.25
	1	14.75	15.5	16.25	17	17.75	18.5	19.25	20	20.75	21.5	22.25	23	23.75	24.5	25.25	26
	2	14.5	15.25	16	16.75	17.5	18.25	19	19.75	20.5	21.25	22	22.75	23.5	24.25	25	25.75
	3	14.25	15	15.75	16.5	17.25	18	18.75	19.5	20.25	21	21.75	22.5	23.25	24	24.75	25.5
	4	14	14.75	15.5	16.25	17	17.75	18.5	19.25	20	20.75	21.5	22.25	23	23.75	24.5	25.25
	5	13.75	14.5	15.25	16	16.75	17.5	18.25	19	19.75	20.5	21.25	22	22.75	23.5	24.25	25
	6	13.5	14.25	15	15.75	16.5	17.25	18	18.75	19.5	20.25	21	21.75	22.5	23.25	24	24.75
	7	13.25	14	14.75	15.5	16.25	17	17.75	18.5	19.25	20	20.75	21.5	22.25	23	23.75	24.5
	8	13	13.75	14.5	15.25	16	16.75	17.5	18.25	19	19.75	20.5	21.25	22	22.75	23.5	24.25
	9	12.75	13.5	14.25	15	15.75	16.5	17.25	18	18.75	19.5	20.25	21	21.75	22.5	23.25	24
	10	12.5	13.25	14	14.75	15.5	16.25	17	17.75	18.5	19.25	20	20.75	21.5	22.25	23	23.75
	11	12.25	13	13.75	14.5	15.25	16	16.75	17.5	18.25	19	19.75	20.5	21.25	22	22.75	23.5
	12	12	12.75	13.5	14.25	15	15.75	16.5	17.25	18	18.75	19.5	20.25	21	21.75	22.5	23.25
	13	11.75	12.5	13.25	14	14.75	15.5	16.25	17	17.75	18.5	19.25	20	20.75	21.5	22.25	23
	14	11.5	12.25	13	13.75	14.5	15.25	16	16.75	17.5	18.25	19	19.75	20.5	21.25	22	22.75
15	11.25	12	12.75	13.5	14.25	15	15.75	16.5	17.25	18	18.75	19.5	20.25	21	21.75	22.5	

## WHAT HAPPENS IN EACH ROUND

At the beginning of each round (except for the first one), the result of the random draw that took place in **the previous period** is recalled to each participant. Each participant is informed that:

- either his group remains the same as in the previous round;
- *BC*: or in the ongoing round he will play with a different person. In this case, each participant is also informed about the level of contribution of his current group member in the previous period.
- *EC*: or in the ongoing round he will play with a different person. In this case, each participant is also informed about the level of contribution of his current group member and the number of points he received in the previous period.

Every round contains 3 stages:

**Stage 1.** Each participant chooses his level of contribution to the pool.

**Stage 2.** Each participant observes the outcomes of a random draw that determines his group in the next round.

**Stage 3.** *BC*: Finally, every participant is informed about his group member's level of contribution and his own gain for the round.

*EC*:

- At the beginning of this stage, every participant is informed about his group member's level of contribution and his own gain for the round.
- Then, each participant has to possibility to express his opinion about the other group member by assigning him a certain number of points. **A high number of points expresses disapproval: 10 points correspond to the strongest disapproval, 0 points correspond to the weakest disapproval.** To do this, you should select the number of points on your screen and press OK button to confirm your choice. The choice of the number of points **has no impact on either participants' gains in the experiment.**
- Finally, **every participant is informed about the number that were assigned to him** by the other group member.

At the end of each round, a message on your screen will inform you that either a new round is about to start, or that the experiment ends.

## PAYMENT OF YOUR EARNINGS

At the end of the experiment, **one round is selected at random**. Each participant receives a sum in EUR corresponding to the amount he earned in this round, converted from ECU to EUR using an exchange rate  $1 \text{ ECU} = 0.40 \text{ EUR}$ , plus a bonus of 5 € for completing the experiment. All payments are made individually and in cash.



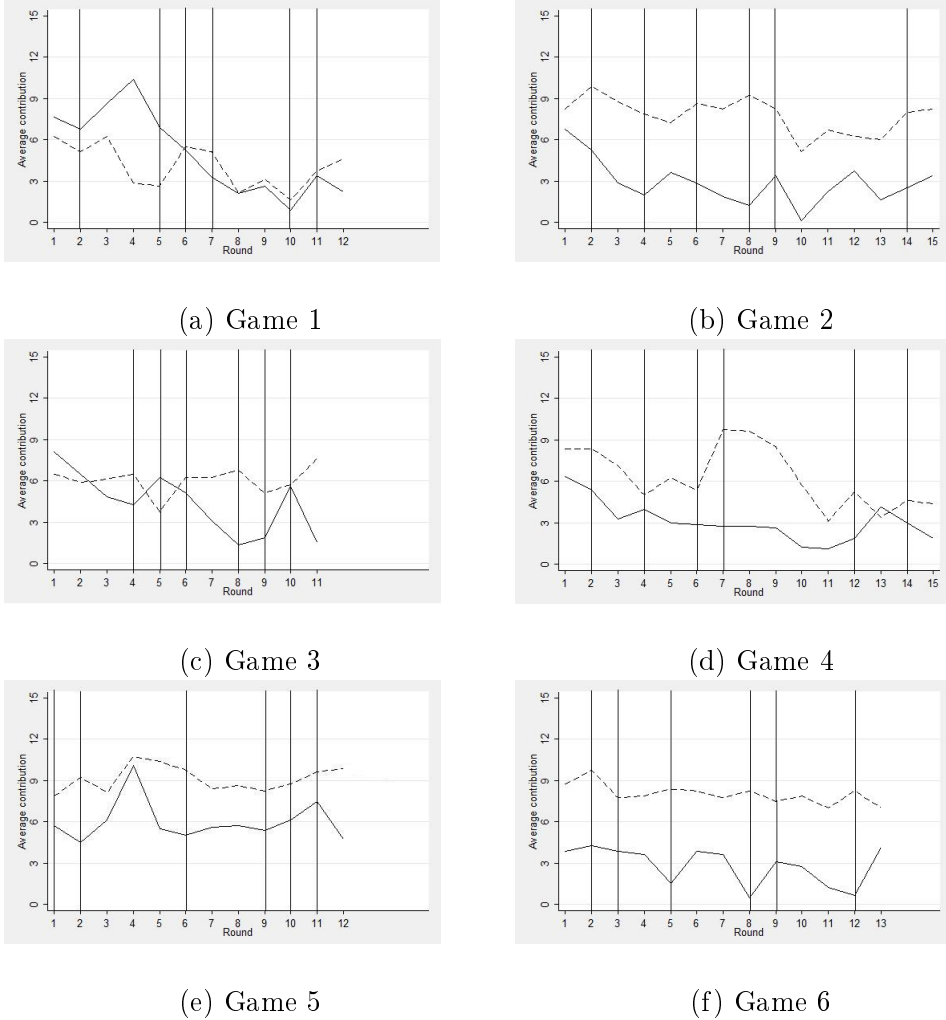
For obvious reasons, **you are not allowed to talk during the experiment.** Participants who violate this rule will be excluded from the experiment and all payments. It is crucial that you understand perfectly the rules of this experiment. Should you have any questions to ask, please raise your hand.

**Thank you for your participation!**

## Evolution of contributions in experimental games

Figures 4a-f suggest that treatment effect is stable over time: in Games 2, 5, and 6 contributions in communication condition dominate those in baseline condition in all rounds; in Games 3 and 4 this tendency is reversed in only 3 rounds out of 11, and 1 round out of 15, respectively. Solely in Game 1 the relationship is more ambiguous.

Figure 4: Evolution of average contribution by game and treatment



**Note.** Vertical lines indicate rounds after which groups are being re-matched. Solid line indicates baseline condition, dashed line indicates communication condition.

## Robustness check: OLS vs. double-censored tobit regressions

This section provides a comparison of different specifications (OLS vs. double-censored tobit) of regression models presented in Tables 2-4. For each model, the columns contain the coefficients from an OLS regression, as well as the coefficients and average marginal effects (ME) from a double-censored tobit regression. \*/\*\*/\*\* indicate statistical significance at the 10%/5%/1% level. Altogether, this evidence suggests that both estimation methods yield similar results, with OLS being more conservative in the assessment of treatment effects and therefore used in the paper.

Table 6: Determinants of group behavior: extension of Table 2

Dep. variable:	Model 1 $c_t^i + c_t^j$			Model 2 $ c_t^i - c_t^j $		
	$\beta_{OLS}$	$\beta_{Tobit}$	$ME_{Tobit}$	$\beta_{OLS}$	$\beta_{Tobit}$	$ME_{Tobit}$
<i>Intercept</i> ( $\beta_0$ )	8.032	6.741	—	4.422*	3.534	—
$1_{Re-matching_{t-1}^t}$ ( $\beta_1$ )	0.753	1.996*	1.365*	0.928*	1.659*	1.184**
$1_{CC}$ ( $\beta_2$ )	6.375*	7.652*	6.177**	0.838	1.229	0.877
$1_{CC} \times 1_{Re-matching_{t-1}^t}$ ( $\beta_3$ )	-0.188	-1.013	-0.816	-0.785	-1.311	-0.935
Additional controls:						
<i>Round</i>						
3	-1.627	-2.211*	-1.873**	1.580	1.979	1.467
4	-0.938	-1.208	-1.029	0.938	0.911	0.659
5	-2.803**	-3.171**	-2.670**	0.551	0.537	0.385
6	-2.241**	-2.740**	-2.313*	0.551	0.708	0.510
7	-2.960**	-3.517**	-2.952***	0.080	0.061	0.043
8	-3.860**	-4.883**	-4.050***	-0.839	-1.359	-0.916
9	-3.741***	-4.788***	-3.975***	0.592	0.469	0.335
10	-5.205***	-6.339***	-5.175***	-0.976	-1.614	-1.077
11	-4.553**	-5.694**	-4.683***	0.488	0.434	0.310
12	-4.317***	-5.594***	-4.606***	0.704	0.602	0.433
13	-4.587**	-5.693**	-4.683***	0.687	0.508	0.364
14	-3.550**	-5.123**	-4.241***	1.666	1.326	0.969
15	-4.334**	-5.829**	-4.787***	0.630	0.501	0.359
<i>Game :</i>						
2	2.067	2.569	2.019	-0.616	-0.688	-0.504
3	1.228	1.624	1.258	-1.408	-1.603	-1.146
4	0.719	1.421	1.098	-1.268	-1.139	-0.825
5	6.636	8.013	6.633	-0.386	-0.328	-0.242
6	2.249	2.789	2.198	-1.263	-1.265	-0.913
<i>N</i>	576					

Table 7: Contributions and historical information in newly formed pairs: extension of Table 3

Determinants:	$\theta_{OLS}$	$\theta_{Tobit}$	$ME_{Tobit}$	$\theta_{OLS}$	$\theta_{Tobit}$	$ME_{Tobit}$	$\theta_{OLS}$	$\theta_{Tobit}$	$ME_{Tobit}$
<i>Intercept</i> ( $\theta_0$ )	4.133***	2.320**	—	7.200***	6.976***	—	4.304**	2.932	—
<i>Group member's</i> :									
<i>Contribution</i> <sub><math>t-1</math></sub> ( $\theta_1$ )	0.228***	0.344***	0.201***	—	—	—	0.291**	0.408*	0.260**
<i>Received_points</i> <sub><math>t-1</math></sub> ( $\theta_2$ )	—	—	—	-0.080	-0.140	-0.089	0.158	0.193	0.123
<i>N</i>	288								

Table 8: Sent disapproval points and the relative size of contribution: extension of Table 4

Determinants:	$\gamma_{OLS}$	$\gamma_{Tobit}$	$ME_{Tobit}$
<i>Intercept</i> ( $\gamma_0$ )	6.227***	7.831***	—
<i>Group_member's_contribution</i> <sub><math>t</math></sub> ( $\gamma_1$ )	-0.380***	-0.766***	-0.433***
<i>Sender's_positive_deviation</i> <sub><math>t</math></sub> ( $\gamma_2$ )	0.167***	0.314***	0.182***
<i>Sender's_absolute_negative_deviation</i> <sub><math>t</math></sub> ( $\gamma_3$ )	0.075	0.128	0.074
$1_{Re-matching_t^{t+1}}$ ( $\gamma_4$ )	-0.536	-1.113	-0.644
$1_{Re-matching_t^{t+1}} \times \text{Group\_members's\_contribution}_t$ ( $\gamma_5$ )	0.047	0.066	0.038
$1_{Re-matching_t^{t+1}} \times \text{Sender's\_positive\_deviation}_t$ ( $\gamma_6$ )	0.010	0.048	0.028
$1_{Re-matching_t^{t+1}} \times \text{Sender's\_absolute\_negative\_deviation}_t$ ( $\gamma_7$ )	-0.138	-0.220	-0.127
<i>N</i>	576		